Review

On epistasis: why it is unimportant in polygenic directional selection

James F. Crow*

Genetics Laboratory, University of Wisconsin, Madison, WI 53706, USA

There is a difference in viewpoint of developmental and evo-devo geneticists versus breeders and students of quantitative evolution. The former are interested in understanding the developmental process; the emphasis is on identifying genes and studying their action and interaction. Typically, the genes have individually large effects and usually show substantial dominance and epistasis. The latter group are interested in quantitative phenotypes rather than individual genes. Quantitative traits are typically determined by many genes, usually with little dominance or epistasis. Furthermore, epistatic variance has minimum effect, since the selected population soon arrives at a state in which the rate of change is given by the additive variance or covariance. Thus, the breeder's custom of ignoring epistasis usually gives a more accurate prediction than if epistatic variance were included in the formulae.

Keywords: quantitative inheritance; epistasis; dominance

1. INTRODUCTION

In its original introduction by Bateson, the word *epistasis* was defined as the masking of the effect of an allele by another at a different locus. It was soon discovered that epistasis is often not complete, and the definition was extended to include partial masking. Fisher further extended the usage to a second definition. He coined the word *epistacy* to mean any departure from additivity—or on a different scale, multiplicativity—of allelic effects. (For small differences additive and multiplicative systems are much the same; I shall ignore any distinction and use additivity from here on.) Fisher's word did not catch on—users preferred epistasis—but the concept did, and his paper marked the beginning of the quantitative analysis of selection. For a discussion of the various meanings of the word, see Phillips (1998). In my view, retaining Fisher's word might have averted some confusion.

In this essay, I want to emphasize this difference in usage. Students of development or evo-devo usually employ the first definition. Animal and plant breeders, or those interested in prediction of evolutionary change, usually use the second. Each is correct in context. My main objective here is to show that the breeders' practice of ignoring epistasis in quantitative selection is fully justified. Much of the material here is taken from an earlier paper (Crow 2008).

It is a pleasure to write this article in honour of Brian Charlesworth, whose contributions to both theoretical and experimental genetics—including

quantitative traits and epistasis—have greatly enriched the field.

2. CONTRASTING VIEWS

Recent years have seen an increased emphasis on epistasis (e.g. Wolf *et al.* 2000; Carlborg & Haley 2004). Students of development and evo-devo, as well as some human geneticists, have paid particular interest to interactions. For those in these fields, epistasis is an interesting phenomenon on its own and studying it gives deeper insights into developmental and evolutionary processes. Ultimately one wants to know which individual genes are involved, and if one is studying the effects of such genes, it is natural to consider the ways in which they interact. Historically, among many other uses, epistasis has provided a means for identifying steps in biochemical and developmental sequences. More generally, including epistasis is part of the description of gene effects. So epistasis, despite methodological challenges, is usually welcomed as providing further insights. Students of development or evo-devo typically study genes of major effect. Of course, genes with major effects are more easily discovered, so they may be providing a biased sample. But we can say that at least some of the genes involved have large effects. And such genes typically show considerable dominance and epistasis.

In contrast, animal and plant breeders have traditionally regarded epistasis as a nuisance, akin to noise in impeding or obscuring the progress of selection. It may seem surprising that the traditional practice of ignoring epistasis has not led to errors in prediction equations. Why? It is this seeming paradox that I wish to discuss.

Continuously distributed quantitative traits typically depend on a large number of factors, each

*jfcrow@wisc.edu

One contribution of 16 to a Theme Issue 'The population genetics of mutations: good, bad and indifferent' dedicated to Brian Charlesworth on his 65th birthday.

making a small contribution to the quantitative measurement. In general, the smaller the effects, the more nearly additive they are. Experimental evidence for this is abundant. This is expected for reasons analogous to those for which taking only the first term of a Taylor series provides a good estimate. For the partial or complete absence of dominance in genes with small viability effects, see Greenberg & Crow (1960); for the minimum influence of epistasis, see Temin *et al.* (1969). (Greenberg and Temin are the same person, before and after marriage.) Subsequent work in many laboratories has abundantly confirmed these conclusions. I conclude that most quantitative traits involve many genes with little dominance or epistasis.

3. POLYGENIC TRAITS

Historically, human height has been regarded as a paradigm of a quantitative trait. The data were consistent with a large number of genes acting in roughly additive fashion (Fisher 1918). Recently, three genome-wide association tests have documented the large number of loci involved. The three studies identified a total of 54 loci (Visscher 2008). Since there was almost no overlap in the three studies, the great majority of loci must have not yet been identified. These 54 loci accounted for about 9 per cent of the genetic variance; hence the total number of loci must be roughly $54 \times (100/9) = 600$. This is a minimum estimate, since only those loci contributing at least 0.3 per cent of the variance would have been detected. So, clearly, human height fits the picture of a trait determined by a large number of genes, each with a very small effect.

Studies of quantitative trait loci (QTLs) in *Drosophila* and mice are roughly concordant with the human data (Flint & Mackay 2009). Rather than a sharp difference between small and large effects, there is a continuum. The distribution is roughly exponential with an accumulation of alleles with the smallest effect. There is often dominance and epistasis in these studies, but in order to be measured by QTL methods, an allele has to have a minimum effect; hence interactions are not surprising. The precise correlation between size of effect and degree of non-additivity is yet to be measured. We should expect that, for quantitative traits, although much of the genetic variance comes from alleles with very minor effects, there are occasionally larger ones.

To account for the very high deleterious human mutation rate without incurring a tremendous genetic load, it is customary to invoke epistasis. I believe that the most effective epistasis is not a consequence of gene action, but rather of the way selection operates. With truncation selection, long known by breeders to be the most efficient method, individuals with a number of mutations above a threshold are eliminated. Thus, harmful mutations are eliminated in bunches. This is what I call quasi-epistasis, generated by selection's grouping alleles of similar effect. Thus, even genes with very small effects are effectively highly epistatic, despite being physiologically additive. It is important to note that truncation does not have to be sharp; approximate rank-order selection is almost

as effective. Although strict truncation in nature is unlikely, quasi-truncation is expected in resource-limited species, and that is a lot of species. For a discussion see Crow (2008) and references therein.

The most extensive selection experiment, at least the one that has continued for the longest time, is the selection for oil and protein content in maize (Dudley 2007). These experiments began near the end of the nineteenth century and still continue; there are now more than 100 generations of selection. Remarkably, selection for high oil content and similarly, but less strikingly, selection for high protein, continue to make progress. There seems to be no diminishing of selectable variance in the population. The effect of selection is enormous: the difference in oil content between the high and low selected strains is some 32 times the original standard deviation. This is all the more remarkable, in that only 12 years were selected each generation. Since the original variance was not large, this experiment dramatically illustrates the ability of Mendelian populations to contain hidden variability that, when brought out, goes far beyond the existing limits (Crow 1992).

Most geneticists would have predicted that the high-oil population would reach a plateau, but no such thing seems to be happening. There is an approach to a plateau in the strain selected for low oil, but this is an obvious consequence of the barrier at zero. When analysis of variance was performed, the variance was mainly additive with only slight effects of dominance and epistasis. The results appear to be consistent with multiple, additive factors, and this was confirmed by QTL analysis (Laurie *et al.* 2004; Hill 2005).

The human and maize data are typical of quantitative traits. Multiple factors with individually small effects acting in a near-additive manner seem to be the rule (Hill *et al.* 2008).

4. WHY DOES THE SELECTED POPULATION NOT EXHAUST ITS VARIANCE?

It has frequently been pointed out that selection uses up genetic variance; therefore, long continued selection should result in a steady decrease of genetic variance (Robertson 1955). Yet, this does not always happen, as the corn experiments illustrate. There are several possible explanations. To me, the most likely is as follows (Crow 1992). The variance in allele frequency, p , is proportional to pq , where $q = 1 - p$. This is a maximum when $p = 1/2$. Most of the variance is contributed by alleles of intermediate frequency, say, between 1/4 and 3/4. As the favourable allele increases by selection, after it passes 1/2 it will make a decreasing contribution to the selectable variance until this becomes close to zero as the allele approaches fixation. Thus, the population variance decreases under selection, as has often been asserted.

But at the same time rare alleles that are favoured by selection will increase and make ever-increasing contributions to the variance as they become common. Thus, the depletion of the variance by fixation of favoured alleles is compensated by bringing previously rare alleles into the range where they contribute

substantially to the variance. In this way, one can easily understand the maintenance of variance during a long-term increase in selected traits, as in the maize experiments. (I am not sure of the origin of this idea. It is clearly not new. I first became aware of it many years ago in a discussion with the late L. N. Hazel.)

This maintenance of variance was recently emphasized by Barton & Keightley (2002), Barton & Coe (2009) and Barton & de Vlader (2009). They point out that if they use the Fisher transformation, $z = \log(p/q)$, z changes linearly with time under selection. Thus, they expect a rough balance between genes making an increased contribution to the variance and those making a decreased contribution, so that variance stays roughly constant.

I think this conclusion is not dependent on the form of the mutational distribution. All that is required is that there be a substantial supply of rare alleles, many of them perhaps in a mutation–selection balance that was reached before the current selection program started.

Of course, in a long enough time the variance will be exhausted. But this may be a very long time and there is always mutation. The extent to which this is important in maintaining quantitative variance under selection is not clear, but may be substantial (Hill 1982). Other possibilities are discussed in Crow (2008).

5. FITNESS AND TRAITS CORRELATED WITH IT

Robertson (1955) has asserted that the (narrow sense) heritability of fitness should be near zero. This is entirely reasonable, but has been difficult to test experimentally. An implication is that traits highly correlated with fitness should also have low heritability. Many of these are likely to be traits with an intermediate optimum and to be close to an equilibrium, at which state the parent–offspring covariance should be very low.

How do we reconcile this observation with the remarkable selectability of oil and protein content of maize and human height? I suggest that neither of these traits is highly correlated with fitness at present. Hence, the rapid response to selection and the large additive variance.

6. THE QUANTITATIVE GENETICIST'S APPROACH

There is another striking difference in the viewpoint of developmental and quantitative genetics. In developmental genetics and often in evo-devo, the effects of individual genes are the objects of study, along with their interactions. In this way one gets a deeper understanding of the developmental or evolutionary process.

The quantitative geneticist does not observe the effects of individual genes, but rather quantitative phenotypes. The idea is to make predictions from macroscopic measurements, such as means, variances and covariances. It is analogous to classical thermodynamics. To measure temperature, one does not average the kinetic energy of individual molecules, but instead uses a thermometer. The individual details

are not needed. Quantitative genetics uses knowledge of the microscopic behaviour of genes to derive macroscopic formulae, but for prediction of individual selection experiments knowledge of contribution of individual genes is not needed.

Fisher's (1930, p. 35) Fundamental Theorem of Natural Selection says that the rate of change of fitness is given by the genetic (additive) variance of fitness at that time. The genetic variance is the sum of the squared deviations of the least squares estimates of the individual genes. This is determined, however, not by measurement of individual genes but by population values such as means, variances and covariances.

Several quantitative geneticists have pointed out analogies with statistical mechanics. To understand the macroscopic properties of the population one does not need to know the details of the microscopic gene behaviour. A particularly enlightening analysis has been recently presented by Barton & de Vlader (2009). They develop a theory of dynamic changes on the assumption that a certain measure of entropy is maximized.

Fisher (1930) was thinking along similar lines when he analogized his Fundamental Theorem of Natural Selection with thermodynamics. In his view, fitness had some properties similar to those of entropy.

In traditional genetic analysis the data consist of means, variances and covariances. In particular, analysis of variance separates additive from dominance and epistatic components of the variance. These can be estimated by various methods, usually involving covariances between relatives. In particular, the dominance component does not contribute to parent–offspring correlation and hence does not contribute to mass selection. Similarly, those epistatic components that depend on the interaction of dominance components do not either; only combinations of additive components contribute. The upshot is that, although there may be large dominance and epistatic components, selection acts only on the additive variance and what is usually a small part of the epistatic variance (Cockerham 1954; Crow & Kimura 2009, p. 132 ff).

For these reasons, one would expect that epistatic variance would have only a small effect on predicting the progress of selection. But that is not all. I'll now consider Thomas Nagylaki's tour de force.

7. NAGYLAKI'S VERY GENERAL TREATMENT

Following several increasingly general studies, Nagylaki (1993) was able to develop the theory under very general conditions. He assumed a discrete-generation, monoecious diploid population mating at random, the usual assumption. But importantly, the number of loci, linkage map, dominance and epistasis are arbitrary. No one had previously treated such a realistic model. The genotypic frequencies might depend on time and on gametic frequencies. The important results hold under weak selection, where the selection coefficient, s (defined as the relative selective difference between the most and least fit genotype) is small relative to the smallest two-locus recombination frequency, c . Most quantitative traits fit this pattern.

After a short time period, approximately $(\ln s)/\ln(1 - c)$, the population evolves approximately as if in linkage equilibrium. After twice this time interval, the linkage disequilibria are nearly constant. Then Fisher's Fundamental Theorem holds to order s^2 ; that is to say the relative change in fitness per generation is given with remarkable accuracy by the additive genetic variance. For a quantitative character correlated with fitness, the variance is replaced by the covariance of the character and fitness, and the accuracy is of order s .

These approximations hold through most of the gene frequency change. The absolute error is small, although the relative error may be large, in particular when the selection process is nearly complete or is close to an equilibrium.

This is a rather loose summary of Nagylaki's findings and the reader is referred to his paper for more rigorous statements. One nicety that I have omitted is that the covariance is between the *effect* on the character and the *excess* of fitness, to use Fisher's terms (1930, p. 30; 1941). The great generality of this theory is indeed remarkable.

This result was foreshadowed in a more restricted study by Kimura (1965), one effect of which was to stimulate Nagylaki to address the problem.

What this means is that the change of fitness or of a character correlated with fitness is, after a short time and through most of the period of gene frequency change, given by the additive genetic variance or covariance. For a two-locus numerical example, see Crow & Kimura (2009, p. 222).

8. CONCLUSION

Students of development, evo-devo and human genetics often place great emphasis on epistasis. Usually they are identifying individual genes, and naturally the interactions among these are of the very essence of understanding. The individual gene effects are usually large enough for considerable epistasis to be expected.

Quantitative genetics has a contrasting view. The foregoing analysis shows that, under typical conditions, the rate of change under selection is given by the additive genetic variance or covariance. Any attempt to include epistatic terms in prediction formulae is likely to do more harm than good. Animal and plant breeders who ignored epistasis, for whatever reasons, good or bad, were nevertheless on the right track. And prediction formulae based on simple heritability measurements are appropriate.

The power of using microscopic knowledge (genes) to develop macroscopic theory (phenotypes), whereby phenotypic measurements are used to develop prediction formulae, is beautifully illustrated by quantitative genetics theory.

I should like to acknowledge the help that I received from Thomas Nagylaki. This has substantially diminished my confusion about a number of theoretical concepts. I have also profited from occasional contacts with Bill Hill and Brian Charlesworth. I wish there were more, but the Atlantic Ocean is a superb isolating mechanism.

REFERENCES

- Barton, N. H. & Coe, J. 2009 On the application of statistical physics to evolutionary biology. *J. Theor. Biol.* **259**, 317–324. (doi:10.1016/j.jtbi.2009.03.019)
- Barton, N. H. & de Vlader, H. P. 2009 Statistical mechanics and the evolution of polygenic quantitative traits. *Genetics* **181**, 997–1011. (doi:10.1534/genetics.108.099309)
- Barton, N. H. & Keightley, P. D. 2002 Understanding quantitative genetic variation. *Nat. Rev. Genet.* **3**, 11–21. (doi:10.1038/nrg700)
- Carlberg, O. & Haley, C. 2004 Epistasis: too often neglected in complex trait studies. *Nat. Rev. Genet.* **5**, 618–625. (doi:10.1038/nrg1407)
- Cockerham, C. C. 1954 An extension of the concept of partitioning hereditary variance for analysis of covariance among relatives when epistasis is present. *Genetics* **39**, 859–882.
- Crow, J. F. 1992 An advantage of sexual reproduction in a rapidly changing environment. *J. Hered.* **83**, 169–173.
- Crow, J. F. 2008 Maintaining evolvability. *J. Genet.* **87**, 349–353. (doi:10.1007/s12041-008-0057-8)
- Crow, J. F. & Kimura, M. 2009 *An introduction to population genetics theory*. Caldwell, NJ: Blackburn Press.
- Dudley, J. W. 2007 From means to QTL: the Illinois long-term selection experiment as a case study in quantitative genetics. *Crop Sci.* **47**, S20–S31. (doi:10.2135/cropsci2007.04.0003IPBS)
- Fisher, R. A. 1918 The correlation between relatives on the supposition of Mendelian inheritance. *Proc. R. Soc. Edinb.* **52**, 399–433.
- Fisher, R. A. 1930 *The genetical theory of natural selection*, 1999 Variorum edn. Oxford, UK: Clarendon Press.
- Fisher, R. A. 1941 Average excess and average effect of a gene substitution. *Ann. Eugen.* **11**, 53–63.
- Flint, C. & Mackay, T. F. C. 2009 Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Res.* **19**, 723–733. (doi:10.1101/gr.086660.108)
- Greenberg, R. & Crow, J. F. 1960 A comparison of lethal and detrimental chromosomes from natural populations. *Genetics* **45**, 1153–1168.
- Hill, W. G. 1982 Predictions of response to artificial selection from new mutations. *Genet. Res.* **40**, 256–278. (doi:10.1017/S0016672300019145)
- Hill, W. G. 2005 A century of corn selection. *Science* **307**, 683–684. (doi:10.1126/science.1105459)
- Hill, W. G., Goddard, M. E. & Visscher, P. M. 2008 Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet.* **4**, e1000008. (doi:10.1371/journal.pgen.1000008)
- Kimura, M. 1965 Attainment of quasi linkage equilibrium when gene frequencies are changing by natural selection. *Genetics* **52**, 875–890.
- Laurie, C. C. *et al.* 2004 The genetic architecture of response to long-term artificial selection for oil concentration in the maize kernel. *Genetics* **168**, 2141–2155. (doi:10.1534/genetics.104.029686)
- Nagylaki, T. 1993 The evolution of multilocus systems under weak selection. *Genetics* **134**, 627–647.
- Phillips, P. C. 1998 The language of gene interaction. *Genetics* **149**, 1167–1171.
- Robertson, A. 1955 Selection in animals: synthesis. *Cold Spring Harbor Symp. Quant. Biol.* **20**, 225–229.
- Temin, R. G. *et al.* 1969 The effect of epistasis on homozygous viability depression in *Drosophila melanogaster*. *Genetics* **61**, 497–519.
- Visscher, P. M. 2008 Sizing up human height variation. *Nat. Genet.* **40**, C489–C490. (doi:10.1038/ng0508-489)
- Wolf, J. B., Brodie, E. D. & Wade, M. J. (eds) 2000 *Epistasis and the evolutionary process*. Oxford, UK: Oxford University Press.